AF
27W

Please type a plus sign (+) inside this box → [+]

PTO/SB/21 (08-00)
Approved for use through 10/31/2002. OMB 0651-0031
U.S. Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

TRANSMITTAL FORM

(to be used for all correspondence after initial filing)

Application Number	09/935,610
Filing Date	August 24, 2001
First Named Inventor	Huiping LI et al.
Group Art Unit	2672
Examiner Name	Jeffrey A. Brier
Total Number of Pages in This Submission: 27	**Attorney Docket Number** 37112-173148

ENCLOSURES *(check all that apply)*

- [x] Fee Transmittal Form
  - [ ] Fee Attached
- [ ] Amendment / Response
  - [ ] After Final
  - [ ] Affidavits/declaration(s)
- [ ] Extension of Time Request
- [ ] Express Abandonment Request
- [ ] Information Disclosure Statement
- [ ] Certified Copy of Priority Document(s)
- [ ] Response to Missing Parts/ Incomplete Application
  - [ ] Response to Missing Parts under 37 CFR 1.52 or 1.53
- [ ] Assignment Papers *(for an Application)*
- [ ] Drawing(s)
- [ ] Licensing-related Papers
- [ ] Petition
- [ ] Petition to Convert to a Provisional Application
- [ ] Power of Attorney, Revocation Change of Correspondence Address
- [ ] Terminal Disclaimer
- [ ] Request for Refund
- [ ] CD, Number of CD(s) ____
- [ ] After Allowance Communication to Group
- [ ] Appeal Communication to Board of Appeals and Interferences
- [x] Appeal Communication to Group *(Appeal Notice, Brief, Reply Brief)*
- [ ] Proprietary Information
- [ ] Status Letter
- [ ] Other Enclosure(s) *(please identify below)*:

Remarks

SIGNATURE OF APPLICANT, ATTORNEY, OR AGENT

Firm or Individual name		26694 PATENT TRADEMARK OFFICE
Signature		Jeffrey W. Gluck, Ph.D., Reg. No. 44,457
Date	March 31, 2006	

CERTIFICATE OF MAILING

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Commissioner for Patents, P. O. Box 1450, Alexandria, VA 22313-1450b on this date:

Typed or printed name			
Signature		Date	

DOC#737874

VENABLE ATTORNEYS AT LAW

SEND TO: Commissioner for Patents, P. O. Box 1450, Alexandria, 22313-1450.

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

In re application of:	Art Unit: 2672
Huiping LI et al.	
Appl. No. 09/935,610 Confirmation No. 8980	Examiner: Jeffery A. Brier
Filed: August 24, 2001	Atty. Docket No. 37112-173148
For: EXTRACTION OF TEXTUAL AND GRAPHIC OVERLAYS FROM VIDEO	Customer No. 26694 PATENT TRADEMARK OFFICE

BRIEF ON APPEAL

Mail Stop Appeal Brief - Patents
Hon. Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450

Sir:

This is an Appeal to the Board of Patent Appeals and Interferences from the decision of the Examiner found in the Final Office Action of October 24, 2005. A Notice of Appeal was timely filed on February 24, 2006.

(1) REAL PARTY IN INTEREST

The real party in interest in the present appeal is ObjectVideo, Inc.

(2) RELATED APPEALS AND INTERFERENCES

There are no related appeals or interferences, to the knowledge and belief of the undersigned.

(3) STATUS OF CLAIMS

The claims pending in this application are Claims 1-38. In the Final Office Action, Claims 29-34 were allowed, Claims 4-21 and 28 were objected to as depending from a rejected base claim, and Claims 1-3, 22-27, and 35-38 were finally rejected. As noted below, Claims 4-21 and 28 have been amended to place them in condition for allowance. Therefore, Claims 1-3, 22-27, and 35-38 are the claims on appeal, and Claims 4-21 and 28-34 are not on appeal.

(4) STATUS OF AMENDMENTS

Three amendments were filed in response to the Final Office Action. The first Amendment was filed on December 7, 2005, and this Amendment resulted in the issuance of the Advisory Action mailed December 21, 2005, in which it was indicated that the amendments and arguments therein did not place the application in condition for allowance, but that the Amendments would be entered upon appeal. The second Amendment was filed on February 9, 2006, and this resulted in the Advisory Action mailed on February 27, 2006, in which it was indicated that the proposed amendments raise new issues for consideration, and thus would not be entered. Finally, Applicants filed a third Amendment on March 8, 2006 to amend Claim 4 to be independent (and thus render Claims 4-21 and 28 allowable) and to set forth a definitive listing of claims for appeal. This Amendment initially resulted in the Advisory Action dated March 20, 2006, indicating that the proposed amendments raise new issues for consideration and that the proposed amendments would not be entered. However, after further discussion between Applicants' undersigned representative and the Examiner, a further Advisory Action, mailed on March 28, 2006, was issued, indicating that the proposed amendments would be entered for purposes of appeal. Consequently, the claims on appeal are as listed in the listing of claims

found in the Amendment filed on March 8, 2006, which are those listed in Appendix I of this Appeal Brief).

(5) SUMMARY OF THE INVENTION

The claimed invention is directed to the extraction of textual and/or graphical overlays from video. In particular, the claimed invention is concerned with the extraction of static overlays that were previously added to (overlaid on) the video. This may be done, for example, during post-production editing, as explained at paragraph [0002] of the Specification. Such overlays may consist of text and/or graphics.

As shown in Fig. 1, the claimed process involves two steps. The first step involves detection 1 of candidate overlays. This is followed by a second step of verification 2, in which the inventive process verifies that a candidate overlay is an actual overlay. These two steps may be found in both of the independent claims on appeal (Claims 1 and 35). Additionally, as shown in Fig. 1, a third step, post-processing 3, may also be employed, which may be used to refine the extracted blocks corresponding to actual overlays. These steps are generally discussed in the Specification at paragraph [0033].

Detection 1 may be accomplished, for example, using the processing shown in Fig. 2 or in Fig. 8. In Fig. 2, video is scanned 11 on a frame-by-frame basis using a small window of pixels, as discussed in paragraph [0034]. Wavelet decomposition 12 may be applied to the video, followed by feature extraction 13 and neural network processing 14. Fig. 10 further shows that the neural network processing 14 may be followed by further processing 15 (e.g., grouping of pixels likely to be classified as text). See paragraphs [0034]-[0035]. In an alternative embodiment, shown in Fig. 8, detection 1 may use a template-matching approach, as described at paragraphs [0058] ff.

Verification 2 may be implemented as shown in Fig. 3 or, alternatively, as shown in Fig. 9 (discussed below). As shown in Fig. 3, verification 2 may comprise spatial verification 21 and temporal verification 22. As shown in Fig. 4, spatial verification 21 may examine various measures of confidence reflecting how confident the process is that the candidate overlay is actually an overlay. In particular, Fig. 4 shows the use of structure confidence 211 and texture confidence 214 and the use of a weighted sum criterion 215, 216 if the structure confidence test 212 is insufficient to determine that the candidate overlay is an actual overlay. This is discussed in the Specification at paragraphs [0046] ff. An embodiment of how structure confidence 211 may be determined is shown in Fig. 5 and discussed in paragraph [0047]. Structure confidence may involve analyzing a particular area of the video frame (specifically, the candidate overlay) to determine if there are recognizable characters that may form an overlay. Paragraph [0047] of the Specification. Texture confidence determination is discussed in the Specification at paragraph [0049] and may involve averaging the numerical outputs of the aforementioned neural network processing over the pixels within an area being analyzed.

An embodiment of temporal verification 22 is shown in Fig. 6 and discussed at paragraphs [0038]-[0045] of the Specification. A purpose of temporal verification 22 is to examine the persistence of a candidate overlay. A static overlay will persist over a number of consecutive video frames and will remain the same in those consecutive video frames, as discussed at paragraph [0038] of the Specification. The embodiment shown in Fig. 6 addresses this by examining pixel behavior over multiple video frames using a mean-square error criterion. Paragraphs [0039] ff. and Fig. 6.

Post-processing 3 may be used, for example, to eliminate pixels that do not form part of the overlay. This may be done via the embodiment shown in Fig. 7 and described at paragraphs [0053]-[0056]. This embodiment is based on the idea that pixels forming a static textual overlay, for example, should have low temporal variances. The embodiment shown in Fig. 7 computes a

variance for each pixel 32 and compares it with a threshold 33 to determine if the pixel is actually part of the overlay.

An alternative embodiment of verification 2 is shown in Fig. 9. The approach shown in Fig. 9 is based on frame-to-frame correlation 21', which may involve computation of a mean-square error measure of correlation between frames. Paragraph [0061]. In this embodiment, the frame-to-frame correlation may be compared to a threshold value 22' to determine if the candidate overlay is an actual overlay. Again, this approach is directed to determining persistence of a candidate overlay. *Id.*

(6) ISSUES

This Appeal involves the following issues for decision by the Board:

(a) Whether Claims 1-3, 22, 23, 26, and 27 are properly rejected under 35 U.S.C. § 102(b) as being anticipated by Chun et al.;

(b) Whether Claims 1 and 22-27 are properly rejected under 35 U.S.C. § 102(a) as being anticipated by Antani et al.; and

(c) Whether Claims 35-38 are properly rejected under 35 U.S.C. § 102(a) as being anticipated by Antani et al.

(7) ARGUMENTS

I. CLAIMS 1-3, 22, 23, 26, AND 27 ARE ALLOWABLE OVER CHUN ET AL. BECAUSE CHUN ET AL. FAILS TO DISCLOSE ALL OF THE ELEMENTS OF THESE CLAIMS.

Of Claims 1-3, 22, 23, 26, and 27, only Claim 1 is an independent claim; Claims 2, 3, 22, 23, 26, and 27 all depend from Claim 1.

Claim 1 is directed to a method of video processing comprising extracting a pre-existing static overlay present in a video sequence. This extracting comprises detecting at least one

potential overlay in the video sequence and verifying that each potential overlay is an actual static overlay that was previously added to an original video sequence to obtain the video sequence.

Structure of Claim 1. While the preamble of Claim 1 recites "a method of video processing," that method must comprise "extracting a pre-existing static overlay." If a reference does not disclose extracting a pre-existing static overlay, it cannot anticipate the claimed method. Furthermore, Claim 1 recites that the extracting comprises (at least) two components: detecting a potential overlay and verifying that the potential overlay is an actual static overlay. Again, a reference that does not disclose both of these components cannot anticipate the claimed method. As the Federal Circuit has stated, "Anticipation under § 102 requires that a single prior art reference disclose each and every limitation of the claimed invention." Moba, B.V. v. Diamond Automation, Inc., 325 F.3d 1306, 1321, 66 USPQ2d 1429, 1439 (Fed. Cir. 2004).

Discussion of Chun et al. Article. The Chun et al. reference discusses the extraction of text from video. **It is irrelevant whether the text is part of the video scene or if it is (part of) an overlay; Chun et al.'s technique will detect the text.** Therefore, an initial observation is that Chun et al. is not directed to the extraction of overlays. As will be discussed below, this results in an essential deficiency when one attempts to read Chun et al. on the method of Claim 1.

The Final Office Action asserts that Chun et al., Section 3.1, "discusses extracting candidate area for text area." Final Office Action at 5. The Final Office Action further asserts that "Section 3.2 discusses verification of candidates of text area." *Id.* The Final Office Action also adds that Section 2 of Chun et al. discusses "character regions having [*sic*] some fixed colors and sizes, and are densely located in the horizontal region, as shown in Fig. 2. The colors and shapes are not regular in the background." *Id.* Applicants note that this discussion in Section 2 of Chun et al. is merely describing the relationship between "character regions" and

"background regions." Chun et al. at II-1127. Applicants further note that the Chun et al. article makes no distinction between character regions that are part of the video and character regions that are overlaid on the video and that the Chun et al. article makes no distinction between background that is moving relative to the characters (which may happen if the characters are part of the video) and background that is static relative to the characters (which may happen in either case, i.e., if the characters are part of the video or overlaid on the video). That is, **Chun et al.'s technique examines whether there are character regions that are, in some way, "different" from any type of "background" regions.** Therefore, the Final Office Action's assertion, based on this, that "text in the actual video will have movement and will likely have a size different then [*sic*] Chun's algorithm text size" and that "Chun recognized the difference between original video with text and text overlayed [*sic*] onto the original video," the Final Office Action is attempting to make a distinction that is not made in Chun et al. *Id.* at 5-6. **Chun et al. has no capability of differentiating between text in the video scene and text overlaid onto the video scene and will, therefore, detect both. That is, while Section 3.2 of Chun et al. mentioned "Verification of candidates of text area," it is exactly what it says: it verifies whether an area contains text, and not whether or not the text is an actual static overlay that was previously added to an original video sequence, as claimed.** Chun et al. at II-1128.

The Final Office Action additionally notes that Section 5 of Chun et al. discusses "caption area." Final Office Action at 6. However, again, Chun et al. makes no distinction between captions that exist in the original video sequence and captions that may be overlaid onto the original video sequence. That is, Chun et al. may detect a text overlay, but it does not determine whether or not it is an overlay on the original video sequence, and there is no need for this to happen because Chun et al. is directed to detecting any text, not just text that has been overlaid on a video sequence. Once again, **Chun et al. does not verify whether the detected text is or is not added to the original video sequence, as claimed.**

In short, the techniques described in Chun et al. are incapable of determining whether detected text is part of the video or has been overlaid onto the video. As such, **Chun et al. lacks any disclosure or suggestion of verifying that a potential overlay is an actual static overlay, as claimed.** For at least these reasons, it is respectfully submitted that Claim 1 is not anticipated by Chun et al. and that neither are its dependent claims listed above, Claims 2, 3, 22, 23, 26, and 27.

II. CLAIMS 1 AND 22-27 ARE ALLOWABLE OVER ANTANI ET AL. BECAUSE ANTANI ET AL. FAILS TO DISCLOSE ALL OF THE ELEMENTS OF THESE CLAIMS.

Of Claims 1 and 22-27, only Claim 1 is an independent claim; Claims 22-27 depend from Claim 1, either directly or indirectly.

As discussed above, Claim 1 is directed to a method of video processing comprising extracting a pre-existing static overlay present in a video sequence. This extracting comprises detecting at least one potential overlay in the video sequence and verifying that each potential overlay is an actual static overlay that was previously added to an original video sequence to obtain the video sequence.

Structure of Claim 1. While the preamble of Claim 1 recites "a method of video processing," that method must comprise "extracting a pre-existing static overlay." If a reference does not disclose extracting a pre-existing static overlay, it cannot anticipate the claimed method. Furthermore, Claim 1 recites that the extracting comprises (at least) two components: detecting a potential overlay and verifying that the potential overlay is an actual static overlay. Again, a reference that does not disclose both of these components cannot anticipate the claimed method. As the Federal Circuit has stated, "Anticipation under § 102 requires that a single prior art

reference disclose each and every limitation of the claimed invention." Moba, B.V., 325 F.3d at 1321, 66 USPQ2d at 1439.

Discussion of Antani et al. Article. The Antani et al. article discloses techniques for extraction of text from video. The Final Office Action notes, "In the [*sic*] section 3, second paragraph at lines 7-11[,] 'artificial caption text' and 'scene text occurring naturally in a video frame' is [*sic*] discussed." Final Office Action at 8. However, the Final Office Action omits the entire sentence from which it quotes, which is significant in that the actual sentence **makes no differentiation between detection of one or detection of the other**. In particular, the sentence from Antani et al. states, "It should be capable of binarizing both artificial caption text as well as scene text occurring naturally in a video frame."

The Final Office Action attempts to read Section 2 of Antani et al. on the verifying portion of Claim 1. Id. at 9. In particular, the Final Office Action makes the following statement:

> Section 2 discusses the localization stage[,] which uses many methods to localize the text. Section 2 discusses using many different localization algorithms whose outputs are fused in the spatio-temporal decision fusion module over multiple frames to verify that potential text is text. Section 2 also discusses a tracking stage[;] this would inherently verify the potential text is actual text.

Id. The Final Office Action continues by discussing portions of Section 4 and the Abstract of Antani et al., as follows: "The Abstract and section 4 discusses [*sic*] the video having temporal information while the overlayed [*sic*] characters have less temporal information[,] and the overlayed [*sic*] characters are contrasted by a changing background. Text in the original video will more likely have movement from frame to frame." Id. at 9-10. However, **Applicants respectfully submit that none of the cited portions of Antani et al., or any other portion of**

Antani et al., discloses or suggests verifying that a potential overlay is an actual overlay, as claimed.

Treating the cited sections in the order in which they occur in Antani et al., the Abstract states that "text often has poor contrast with a changing background. The proposed system applies a variety of methods and takes advantage of the temporal redundancy in video[,] resulting in good text segmentation." Antani et al. at 831. In other words, the Abstract is merely commenting that, when one tries to extract text from video – any text – a lack of contrast with the background video may be detrimental to the segmentation of the text from the video, and therefore, the authors propose using temporal redundancies in video to alleviate this problem. **This does not disclose verifying that a candidate overlay is an actual overlay.**

Continuing to Section 2 of Antani et al., the authors do, indeed, note that "[t]he video text extraction problem is divided into three main tasks – detection, localization, and segmentation." Id. As an initial observation, it is noted that the problem being addressed is "video text extraction," not overlay extraction. As noted in the Final Office Action, in the portion quoted above, this section of Antani et al. proposes the use of a multiple detection/localization algorithms, followed by the use of a spatio-temporal decision fusion module, and possibly an additional tracking stage. Id. However, Section 2 of Antani et al. notes that the goal of these procedures is "robust text detection," and **there is no provision for differentiating between text within the video and text overlaid on the video.** Id. at 832. That is, there is no verification that any detected text is overlay text, as opposed to text occurring in the video.

Section 4 adds no disclosure of verification that text has been overlaid, as opposed to being text that is merely part of the video. Indeed, Figures 1 reinforces the fact that **no such verification is made, so there is no differentiation between the two types of text.** Id. at 833. Figure 1 shows that both logos, which may have been superimposed on a video scene, and text from a sign or newspaper, which is part of the video scene, are segmented by the Antani et al.

techniques. Id. Hence, Antani et al, it may be argued, even *teaches away* from making such a differentiation, in its goal of extracting *all* text from the video, whether overlaid or not (consistent with the Abstract, as discussed above).

In summary, the techniques described in Antani et al. are incapable of determining whether detected text is part of the video or has been overlaid onto the video. As such, **Antani et al. lacks any disclosure or suggestion of verifying that a potential overlay is an actual static overlay, as claimed.** For at least these reasons, it is respectfully submitted that Claim 1 is not anticipated by Chun et al. and that neither are its dependent claims listed above, Claims 22-27.

III. CLAIMS 35-38 ARE ALLOWABLE OVER ANTANI ET AL. BECAUSE ANTANI ET AL. FAILS TO DISCLOSE ALL OF THE ELEMENTS OF THESE CLAIMS.

Of Claims 35-38, only Claim 35 is an independent claim; Claims 36-38 depend from Claim 35, either directly or indirectly.

Claim 35 is directed to a method of video processing comprising extracting a pre-existing static graphical overlay present in a video sequence. This extracting comprises detecting at least one potential overlay in the video sequence, including performing template matching, and verifying that each potential overlay is an actual static overlay that was previously added to an original video sequence to obtain the video sequence, including performing frame-to-frame correlation of a potential overlay.

Structure of Claim 35. While the preamble of Claim 35 recites "a method of video processing," that method must comprise "extracting a pre-existing static graphical overlay." If a reference does not disclose extracting a pre-existing static graphical overlay, it cannot anticipate the claimed method. Furthermore, Claim 35 recites that the extracting comprises (at least) two

components: detecting a potential overlay and verifying that the potential overlay is an actual static overlay. Again, a reference that does not disclose both of these components cannot anticipate the claimed method. The detecting must comprise performing template matching, and the verifying must include performing frame-to-frame correlation of a potential overlay. An anticipating reference must, similarly, disclose these components, as well, and if it does not, it is not an anticipatory reference. As the Federal Circuit has stated, "Anticipation under § 102 requires that a single prior art reference disclose each and every limitation of the claimed invention." Moba, B.V., 325 F.3d at 1321, 66 USPQ2d at 1439.

Discussion of Antani et al. Article. The Antani et al. article discloses techniques for extraction of text from video. It is noted initially that the arguments used above to show that Antani et al. fails to disclose verifying that a potential overlay is an actual overlay are applicable here, as well. For the sake of brevity, these arguments will not be repeated; one is referred to Section II of this brief. For at least those reasons, it is respectfully submitted that Claim 35 and its dependent claims, Claims 36-38, are not anticipated by Antani et al. However, there are additional deficiencies in Antani et al. that provide further reasons why Antani et al. fails to anticipate Claims 35-38.

The Final Office Action cites Section 2 of Antani et al. as disclosing the use of template matching. Final Office Action at 13. In particular, the Final Office Action states the following:

> Section 2 discusses the detection of potential overlay[s] in the detection stage[,] which consists of many different localization algorithms whose outputs are fused in the spatio-temporal decision fusion module over multiple frames. In order to determine if text exists[,] then predefined knowledge of the text is compared with the current image to determine if a match exists. Predefined knowledge of the text is a template.

Id. On the other hand, Section 2 of Antani et al., reproduced here in its entirety, states the following:

> The video text extraction problem is divided into three main tasks – detection, localization, and segmentation. The recognition (OCR) stage is assumed to lie outside our system. The main components are implemented as POSIX threads. The detection/localization stage consists of a battery of methods for localizing text in the frame. Some methods use the MPEG DCT coefficients, while other use the uncompressed frame. Currently, we have included work from Gargi *et al[.]* [2], Chaddha *et al[.]* [4], LeBourgeois [7], and Mitrea and de With [11]. The spatio-temporal decision fusion module aggregates the decisions of the multiple localization algorithms over multiple frames, defining tight bounding regions around text instances. To improve results, the tracking stage can be used to provide additional input to the spatio-temporal decision-fusion stage. The segmentation module contains the methods to binarize a localized text instance resulting from the fusion process, making it suitable for OCR. The system is designed to take advantage of the temporal nature of video and uses the fact that the text data lasts over several frames for providing robust text detection.

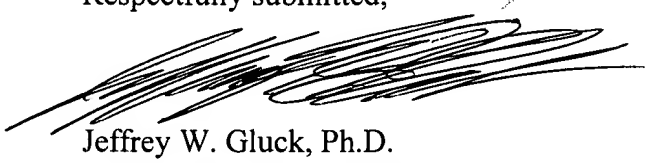
Antani et al. at 831-832. Applicants are unable to find any disclosure of any technique that even resembles template matching in Section 2 of Antani et al. Applicants have also reviewed the rest of Antani et al. and are, similarly, unable to find any such disclosure. It is, for this further reason, respectfully submitted that **Antani et al. fails to disclose template matching**, and therefore, Antani et al. does not anticipate Claim 35 or its dependent claims, Claims 36-38.

Additionally, Applicants note that Claim 35 recites, "extracting a pre-existing *graphical* overlay." Throughout their Specification, Applicants have referred to "text" and "graphics" overlays as different. See, e.g., Specification at paragraphs [0001]-[0004], [0011], [0057], [0062]-[0064]. As discussed above, Antani et al. is limited to detection of *text, only* (and, also as discussed above, not only text overlays). **Antani et al. contains no disclosure of extraction of graphical overlays.** It is, therefore, respectfully submitted that, for this additional reason, Antani et al. does not anticipate Claim 35 or its dependent claims, Claims 36-38.

* * *

The $250.00 fee required by 37 C.F.R. § 41.20(b)(2) for a small entity is hereby authorized to be charged to our Deposit Account No. 22-0261, as are any other fees that may be necessary (any excess payments should be refunded to the same deposit account).

Respectfully submitted,



Date: March 31, 2006

Jeffrey W. Gluck, Ph.D.
(Registration No. 44,457)
VENABLE LLP
P.O. Box 34385
Washington, DC 20043-9998
Telephone: (202) 344-4000
Direct Dial: (202) 344-8017
Facsimile: (202) 344-8300

DC2-731325

(9) APPENDIX I – CLAIMS

The claims on appeal are presented below.

1. A method of video processing to be performed by video processing equipment, the method comprising:

extracting a pre-existing static overlay present in a video sequence, said extracting comprising:

detecting at least one potential overlay in the video sequence; and

verifying that each at least one potential overlay is an actual static overlay that was previously added to an original video sequence to obtain said video sequence.

2. The method of Claim 1, further comprising the step of:

post-processing at least one actual static overlay to remove extraneous pixels.

3. The method of Claim 2, wherein said step of post-processing comprises the steps of:

computing a variance for each pixel of the at least one actual static overlay; and

comparing the variance with a threshold to determine whether or not the pixel should be removed as an extraneous pixel.

4. A method of video processing, to be performed by video processing equipment, comprising:

extracting a pre-existing overlay present in a video sequence, said extracting comprising:

detecting at least one potential overlay in the video sequence; and

verifying that the at least one potential overlay is at least one actual overlay, wherein said step of detecting comprises the steps of:

performing wavelet decomposition on the video sequence;

extracting features based on the results of the wavelet decomposition; and

performing neural network processing on the extracted features.

5. The method of Claim 4, wherein said neural network processing step comprises the step of:

utilizing three-layer back-propagation neural network processing.

6. The method of Claim 4, wherein said step of verifying comprises the steps of:

performing temporal verification; and

performing spatial verification.

7. The method of Claim 6, wherein said step of temporal verification comprises the steps of:

translating said potential overlay over a search range;

for each translated version of said potential overlay, computing a mean square error in a next video frame of said video sequence subsequent to a video frame in which said potential overlay is originally detected;

determining a minimum of the computed mean square errors for said next video frame; and

comparing the determined minimum mean square error to a threshold.

8. The method of Claim 7, further comprising the steps of:

selecting a particular pixel of said potential overlay and recording its coordinates; and

recording the translated coordinates of said particular pixel corresponding to said determined minimum mean square error.

9. The method of Claim 8, further comprising the step of:

if the determined minimum mean square error does not exceed said threshold, determining if the coordinates of said particular pixel of said potential overlay match said translated coordinates of said particular pixel corresponding to said determined minimum mean square error.

10. The method of Claim 9, wherein said determining step determines an approximate match.

11. The method of Claim 9, further comprising the step of:

if said determining step determines that there is not a match, performing the sub-steps of:

incrementing an error count; and

comparing said error count to a predetermined threshold; and

if said determining step determines that there is a match, decreasing said error count.

12. The method of Claim 11, wherein said step of decreasing said error count comprises the step of decrementing said error count.

13. The method of Claim 11, wherein said step of decreasing said error count comprises the step of clearing said error count.

14. The method of Claim 11, wherein said steps of computing, determining, recording, and comparing are performed for subsequent video frames of the video sequence as long as said determined minimum mean square error is found not to exceed said threshold and either the coordinates of said particular pixel of said potential overlay match said translated coordinates of said particular pixel corresponding to said determined minimum mean square error or said error count does not exceed said predetermined threshold.

15. The method of Claim 6, wherein said step of performing spatial verification is performed for a candidate overlay determined by said step of performing temporal verification and comprises the steps of:

determining a structure confidence for said candidate overlay; and

determining a texture confidence for said potential overlay.

16. The method of Claim 15, further comprising the steps of:

determining if said structure confidence meets a first threshold test; and

determining if a weighted sum of said structure confidence and said texture confidence meets a second threshold test.

17. The method of Claim 16, wherein said step of said step of determining a texture confidence is performed only if said structure confidence fails to meet said first threshold test.

18. The method of Claim 17, wherein if either of said steps of determining if said structure confidence or said weighted sum meets said respective first or second threshold test is satisfied for the candidate overlay, the candidate overlay is declared to be an actual static overlay; and wherein said steps of determining if said structure and weighted sum meet said respective first and second threshold tests are not satisfied for the candidate overlay, the candidate overlay is determined not to be an actual static overlay.

19. The method of Claim 15, wherein said step of determining a structure confidence comprises the steps of:

analyzing the candidate overlay to determine characters;

analyzing the determined characters for the presence of words; and

setting a numerical value for said structure confidence depending upon the presence of one or more intact words.

20. The method of Claim 19, wherein said step of setting a numerical value comprises the steps of:

setting the structure confidence equal to one if at least one intact word is detected; and

if no intact word is detected, setting the structure confidence equal to a number of correct characters divided by a total number of characters.

21. The method of Claim 15, wherein said step of determining a texture confidence comprises the step of:

setting the texture confidence equal to an average value of outputs of said neural network processing step corresponding to all the pixels in a potential overlay.

22. The method of Claim 1, wherein said step of detecting comprises the step of:

performing template matching to determine the presence of a potential overlay.

23. The method of Claim 22, wherein said step of detecting further comprises the step of:

determining a template to be used in said step of performing template matching.

24. The method of Claim 22, wherein said step of verifying comprises the steps of:

performing frame-to-frame correlation of said potential overlay; and

comparing a result of the frame-to-frame correlation with a threshold to determine if the potential overlay is an actual static overlay or not.

25. The method of Claim 24, wherein said step of performing frame-to-frame correlation comprises the step of:

forming a mean square error over a set of frames from said video sequence, averaged over all of the pixels in said potential overlay.

26. A computer-readable medium containing computer-executable code for causing a computer to implement the method of Claim 1.

27. A computer system comprising:

a computer; and

a computer-readable medium coupled to said computer and containing computer-executable code for causing said computer to implement the method of Claim 1.

28. A computer system comprising:

a computer;

a computer-readable medium coupled to said computer and containing computer-executable code for causing said computer to implement the method of Claim 4; and

an external processor, in communication with said computer, on which is performed the step of neural network processing.

29. A method of processing video to be performed by video processing equipment, the method comprising:

extracting a pre-existing overlay present in a video sequence, said extracting comprising:

detecting at least one potential overlay in the video sequence, said detecting comprising the steps of:

performing wavelet decomposition on the video sequence;

extracting features based on the results of the wavelet decomposition;

performing neural network processing on the extracted features; and

in parallel with said steps of performing wavelet decomposition, extracting features, and performing neural network processing, performing template matching; and

verifying that the at least one potential overlay is at least one actual overlay.

30. The method of Claim 29, wherein said step of verifying includes the step of:

performing temporal verification.

31. A method of processing video to be performed by video processing equipment, the method comprising:

extracting a pre-existing textual overlay present in a video sequence, said extracting comprising:

detecting at least one potential overlay in the video sequence, said detecting comprising steps of:

performing wavelet decomposition on the video sequence;

extracting features based on the results of the wavelet decomposition; and

performing neural network processing on the extracted features; and

verifying that the at least one potential overlay is at least one actual overlay.

32. The method of Claim 31, wherein said step of verification comprises the steps of:

performing temporal verification; and

performing spatial verification.

33. The method of Claim 32, wherein said step of spatial verification is performed for a candidate overlay output by said step of temporal verification and comprises the steps of:

determining a structure confidence for said candidate overlay;

determining a layout confidence for said candidate overlay; and

determining a texture confidence for said candidate overlay.

34. The method of Claim 32, wherein said step of performing temporal verification comprises the steps of:

computing a mean square error for each pixel of said potential overlay over a set of video frames of said video sequence;

averaging said mean square error for each pixel over all of the pixels in said potential overlay, thus producing an average mean square error; and

comparing said average mean square error to a threshold to determine if the potential overlay is a candidate overlay or not.

35. A method of processing video to be performed by video processing equipment, the method comprising:

extracting a pre-existing static graphical overlay present in a video sequence, said extracting comprising:

detecting at least one potential overlay in the video sequence, said detecting comprising the step of:

performing template matching; and

verifying that each at least one potential overlay is an actual static overlay that was previously added to an original video sequence to obtain said video sequence, said verifying comprising the step of:

performing frame-to-frame correlation of a potential overlay determined by said detecting step.

36. The method of Claim 35, wherein said step of detecting further comprises the step of:

determining a template to be used in said step of performing template matching.

37. The method of Claim 36, wherein said step of determining a template comprises the step of:

performing addition or frame-by-frame subtraction of video frames.

38. The method of Claim 36, wherein said step of determining a template comprises the steps of:

segmenting video frames into foreground and background objects;

performing correlation tracking to determine if any foreground object remains in the same absolute location in each video frame.

APPENDIX II – EVIDENCE

(None)

APPENDIX III – RELATED PROCEEDINGS

(None)